

Methodology

Universe for Statistics of U.S. Businesses

Statistics of U.S. Businesses (SUSB) covers more than 6 million single-unit establishments and 1.8 million multi-unit establishments. The Business Register is the Census Bureau's source of information on employer establishments included in the SUSB program. The Business Register is a multi-relational database that contains a record for each known establishment that is located in the United States with paid employees. An establishment is a single physical location at which business is conducted or services or industrial operations are performed. An establishment is not necessarily equivalent to a company or enterprise, which may consist of one or more establishments. A single-unit enterprise owns or operates only one establishment. A multi-unit enterprise owns or operates two or more establishments. The treatment of establishments on the Business Register differs according to whether the establishment is part of a single-unit or multi-unit enterprise.

Single Unit Enterprise

A single-unit enterprise's unique identifier is its Employer Identification Number (EIN). The Internal Revenue Service (IRS) issues an EIN and the enterprise uses it as an identifier to report payroll taxes. All employers are required to have at least one EIN. Because a single-unit enterprise has only one establishment, there is a one-to-one relationship between the enterprise and the EIN. Thus the enterprise, the EIN, and the establishment all reference the same physical location and all three terms can be used interchangeably and unambiguously when referring to a single-unit enterprise.

Descriptive information for a single-unit establishment in the SUSB universe, including geographic location, industry classification, legal form of organization classification, payroll, and employment, come from a variety of administrative record and survey sources. Administrative records filed by EIN are the most common source of information for single-unit establishments, with updates on geographic location and industry classification from Census Bureau conducted surveys.

Multi-Unit Enterprise

For multi-unit enterprises a different structure connects the enterprise with its establishments via the EIN. Essentially a multi-unit enterprise is associated with a cluster of one or more EINs and EINs are associated with one or more establishments. A multi-unit enterprise consists of at least two establishments. Each enterprise is associated with at least one EIN and only one enterprise can use a given EIN. One multi-unit enterprise may have several EINs. Similarly, there is a one-to-many relationship between EINs and establishments. Each EIN can be associated with many establishments, but each establishment is associated with only one EIN. With the possibility of one-to-many relationships, there must be distinction between the enterprise, its EINs, and its establishments. A unique employer unit identification number identifies each establishment owned by a multi-unit enterprise on the Business Register.

There is less dependency on administrative record sources for multi-unit establishment information because EIN and establishment are not equivalent for multi-unit enterprises. The Census Bureau's Economic Census (conducted every five years in years ending in '2' and '7') initially identifies multi-unit

companies when a company expands to more than one establishment. Establishments for a multi-unit company are identified through the Economic Census and the annual Enterprise Organization Survey (COS). Geographic location, industry classification, payroll and employment come primarily from the Economic Census and the COS. EIN-level administrative payroll and employment data are apportioned to the establishment level in cases of nonresponse or for smaller companies not selected for the COS.

Legal Form of Organization

The Legal Form of Organization (LFO) for SUSB establishments is derived from administrative record sources. SUSB publishes U.S.-level data (starting with 2007) and state-level data (starting with 2018) by the following legal forms of organization:

- C-Corporations and other corporate legal forms of organization
- S-Corporations
- Sole Proprietorships
- Partnerships
- Non-profits
- Government
- Other

LFO data are published by NAICS sector and enterprise employment size class.

Employment Services

Historically, the permanent on-site workforce at a business location were paid employees of that establishment. This traditional practice of enterprises directly hiring employees is still the dominant employer/employee relationship in the United States. However, in recent years a new workforce arrangement has emerged. Under this new arrangement, a professional employer organization (PEO) or employee-leasing enterprise operates in a co-employment relationship with client businesses to provide leased workers and specialized management services. Employee leasing enterprises are responsible for payroll, including withholding and remitting employment-related taxes, for some or all of the employees of their clients, and also serve as the employer of those employees for benefits and related purposes.

SUSB has shown a steady increase in PEOs and other employment services establishments (NAICS 5613), as well as payroll services establishments (NAICS 541214), over the past few years. Sometimes, the employees are not classified in the predominant industry of the client businesses. Employment services establishments may pay these employees out of a single payroll office. This may result in the leasing enterprise's employment and payroll data being reported in the county where the payroll office is located, thus distorting the data for that county. In some cases, many thousands of employees may be paid from a single payroll office. Therefore, for geography purposes, employee-leasing establishments may be published in the "statewide" category in states where such payroll offices are located, as these establishments service multiple counties.

Exclusions and Undercoverage

SUSB covers most of the country's economic activity. The series excludes data on self-employed individuals, employees of private households, railroad employees, agricultural production employees, and most government employees. Businesses operating without an EIN, and businesses with an EIN but without employees, are excluded from the SUSB universe.

A certain amount of undercoverage occurs in the universe, primarily with establishments for multi-unit companies. Usually, the Census Bureau does not create a multi-unit company structure in the Business Register for very small employers (less than 10 employees) identified in the Economic Census. In addition, the COS is an annual mail survey that includes all multi-unit companies with 500 or more employees. Companies with less than 500 employees are only selected for the COS when administrative record sources indicate the company may be undergoing organizational change and are adding or dropping establishments. Establishments for smaller companies may be missed, as well as establishments for companies not responding to the Economic Census or the COS. The Census Bureau takes much effort to get establishment information for large companies because of their importance to the economy. The Census Bureau does not have any estimates of establishment undercoverage. Coverage of payroll and employment is very good because of the usage of administrative record data.

Industry and Geographic Classification

Industry classification of business establishments in SUSB is according to the North American Industry Classification System (NAICS), which includes nearly 1,200 industries. For more information on the NAICS codes, as well as changes over the years, visit the [NAICS website](https://web.archive.org/web/20210901053416/https://www.census.gov/naics/) [https://web.archive.org/web/20210901053416/https://www.census.gov/naics/]. NAICS codes are updated every five years.

The primary source of industry classification is derived from data collected through the Economic Census or through other Census surveys. When this is not available, the Census Bureau uses a hierarchy of administrative record sources to assign a NAICS code, including classifications from the Bureau of Labor Statistics, business birth information, and self-assigned codes from income tax records.

For a small percentage of records, only a partial classification is possible from all sources. For these cases, a complete industry classification is assigned, or imputed, by using a "nearest neighbor" assignment process that assigns the NAICS from an establishment with similar payroll, employment and organizational type with possible constraints. Analysts review the assignments to ensure that anomalies do not occur at the county level. For some multi-unit establishments with a partial classification, a complete code is imputed from another establishment within the same enterprise. The imputation rate for complete codes varies widely during the five-year Economic Census processing cycle, but generally affects small businesses. Completely unclassified records are an even smaller percentage and are tabulated and published separately.

North American Industry Classification System (NAICS)

[<https://web.archive.org/web/20210901053416/https://www.census.gov/naics/>]

The Statistics of U.S. Businesses series includes the following NAICS sectors:

11	Forestry, Fishing and Hunting, and Agricultural Support Services (NAICS 113-115)
21	Mining
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
99	Industries Not Classified

Industries Excluded from Statistics of U.S. Businesses

SUSB covers most NAICS industries excluding crop and animal production (NAICS 111,112), rail transportation (NAICS 482), Postal Service (NAICS 491), pension, health, welfare, and vacation funds (NAICS 525110, 525120, 525190), trusts, estates, and agency accounts (NAICS 525920), office of notaries (NAICS 541120), private households (NAICS 814), and public administration (NAICS 92).

Government Establishments Included in Statistics of U.S. Businesses

Although most government establishments are excluded from tabulation, SUSB includes government sponsored wholesale liquor establishments (NAICS 4248), retail liquor stores (NAICS 44531), tobacco stores (NAICS 453991), book publishers (511130), monetary authorities – central bank (NAICS 521110), federally-chartered savings institutions (NAICS 522120), federally-chartered credit unions (NAICS 522130), hospitals (NAICS 622), gambling industries (NAICS 7132), and casino hotels (NAICS 721120).

Enterprise Classification

An enterprise may have establishments in many different industries. For the purpose of classifying an entire enterprise into a single industry, the classification methodology starts by excluding nonoperating establishments – establishments classified as manufacturers' sales branches and offices, establishments engaged in management of enterprises and enterprises (NAICS 55), and auxiliary establishments. The enterprise is then classified into the 2-digit NAICS sector in which it paid the largest share of its payroll. Then, within this 2-digit NAICS sector, the enterprise is classified into the 3-digit NAICS sub-sector in which the enterprise paid the largest share of payroll. Finally, within the assigned 3-digit NAICS sub-sector, the enterprise is classified into the 4-digit NAICS industry group with the largest share of payroll.

Geographic Classification of Establishments

SUSB classifies an establishment by its physical location. Under the usual definition, an establishment or business is a fixed physical location or permanent structure where some form of business activity is conducted. The Economic Census and the COS requests the physical location of each establishment in a company. In addition, administrative record sources provide physical location addresses. In some cases when the physical location is not available geographic assignment is based on the mailing address. When a business relocates, there may be a significant delay until the Census Bureau receives the updated physical location address, particularly for small businesses. This may have an impact on establishment counts at the county level, but this level is not measured.

SUSB publishes county data for Louisiana (parishes) and Alaska (organized boroughs, city and boroughs, municipalities, and census areas) as the equivalent of a county. The independent cities in Virginia, and the cities of Baltimore, MD; Carson City, NV; and St. Louis, MO, are treated as separate counties.

Employers without a fixed location within a state (or of unknown county location) are included under a “statewide” classification at the end of the county tables. This incomplete detail causes only slight understatement of county employment.

Protecting Confidentiality

In accordance with U.S. Code, Title 13, Section 9, no data are published that would disclose the operations of an individual employer.

The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied.

Noise Infusion

Since reference year 2007, SUSB has used Noise Infusion to protect the confidentiality of respondent data. Noise Infusion is a method of disclosure avoidance in which data values for each establishment are perturbed prior to table creation by applying a random noise multiplier to the magnitude data (i.e., characteristics such as first-quarter payroll, annual payroll, and number of employees) for each establishment. Disclosure protection is accomplished in a manner that results in a relatively small change in the vast majority of cell values. Each published cell value has an associated noise flag indicating the relative amount of distortion in the cell value resulting from the perturbation of the data contributing to the cell. The flag for ‘low noise’ (G) indicates the cell value was changed by less than 2 percent with the application of noise, the flag for ‘moderate noise’ (H) indicates the value was changed by at least 2 percent but less than 5 percent, and the flag for ‘high noise’ (J) indicates the value was changed 5 percent or more. Prior to 2015, cell values with a ‘high noise’ flag, were suppressed from publication by replacing the cell value and associated noise flag with a “D.” For an introduction to the noise infusion disclosure avoidance method, see [Using Noise for Disclosure Limitation of Establishment Tabular Data \[PDF\]](#)

[<https://web.archive.org/web/20210901053416/https://www.census.gov/prod/2/gen/96arc/iaevans.pdf>] by Timothy Evans, Laura Zayatz, and John Slanta in the Journal of Official Statistics (1998). Other cells in the table may be suppressed in the same manner because of concerns about the quality of the data. Though some of these suppressed cells may be derived by subtraction, the results are not official and may differ substantially from the actual, unknown population values.

Prior to reference year 2017, the number of establishments in a particular tabulation cell was not considered sensitive; therefore, counts of establishments were released without any disclosure avoidance methods applied. Beginning with reference year 2017, cells with fewer than 3 establishments have been omitted from the release. The use of an employment range flag (EMPLFL_R) was discontinued beginning in reference year 2018 and a noisy employment cell value is provided.

Also starting with 2017 SUSB, the publication flag 'D' (withheld to avoid disclosing data for individual businesses; data are included in broader industry totals) has been replaced with publication flag 'S' (withheld to avoid releasing data that do not meet publication standards; data are included in broader industry totals).

Prior to the use of noise infusion, SUSB data were protected using cell suppression implemented using the p-percent or (n,k) rule.

Reliability of Data

Payroll and employment data are obtained from administrative records for single-unit enterprises and a combination of administrative records and survey-collected data for multi-unit enterprises. They are not subject to sampling error, but are subject to nonsampling errors, which can be attributed to several sources: inability to identify all cases that should be in the universe; definition and classification difficulties; errors in recording or coding the data obtained; and other errors of coverage, processing, and estimation for missing or misreported data.

The accuracy of these tabulated data is determined by the joint effects of the various nonsampling errors. No direct measurement of these effects has been obtained except for estimation for missing or misreported industry classifications; however, precautionary steps were taken in all phases of the processing to minimize the effects of nonsampling errors.

Employment is either missing or reported as zero, when quarterly payroll is greater than zero, for approximately 6.5% of incoming administrative records. In addition, less than one percent of employment values are reported as positive, but the average wage falls outside of expected limits. For either of these situations, employment is imputed using one of several methods. The most frequent is using the average wage for the industry and geographic area. Other methods include using the enterprise's average employment of the two adjacent quarters, or average enterprise wage data for the prior year. Quarterly payroll is edited by comparing with reported data from other quarters over a two-year period to determine any anomalies and potential misreporting. Suspected missing payroll and extreme values are imputed based on enterprise reporting patterns over the two-year period. The Census Bureau imputes payroll for less than one percent of all incoming administrative payroll records.

Establishment payroll and employment data for multi-unit enterprises are collected through the Economic Census and the COS. Data for companies not included in the COS or not responding to the survey are imputed from administrative record data by taking company level administrative payroll and employment data and distributing it down to the establishment level by best estimates of the size of each establishment in the company. If some establishments have reported payroll and some do not, the breakdown is performed with the difference between the administrative data at the company level and the total reported amounts.

Comparability

The comparability of data over time may be affected by changes in industry classifications, definitions of establishments, establishment's active status, and/or changes to geographic boundaries (actual or statistically-defined areas).

Industry Classification

Since 1998, Statistics of U.S. Businesses (SUSB) has been tabulated based on the North American Industry Classification System (NAICS). For more information on NAICS, see the NAICS information page [<https://web.archive.org/web/20210901053416/https://www.census.gov/naics/>] . Data were tabulated according to the Standard Industrial Classification (SIC) [<https://web.archive.org/web/20210901053416/https://www.osha.gov/pls/imis/sicsearch.html>] System for prior periods.

- 2017 to 2021 data use 2017 NAICS
- 2012 to 2016 data use 2012 NAICS
- 2008 to 2011 data use 2007 NAICS
- 2003 to 2007 data use 2002 NAICS
- 1998 to 2002 data use 1997 NAICS
- 1988 to 1997 data use 1987 SIC

Classification changes are explored in detail in NAICS concordance reference files

[<https://web.archive.org/web/20210901053416/https://www.census.gov/naics/?68967>] .

Prior to 2012, SUSB lagged by one year in the adoption of the classification system employed in the economic census. Starting in 2012, the classification system was changed in the same year.

Geography

Counties: County boundary changes may occur between each publication year; however these changes are implemented in batch once every 5 years. Such changes are detailed in the following page:

Substantial Changes to Counties and County Equivalent Entities: 1970-Present

[<https://web.archive.org/web/20210901053416/https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html>] . This is primarily done to maintain data consistency and comparability over time.

Metropolitan/Micropolitan Areas (MSAs): MSAs are redefined after each decennial census and may be revised based on population estimates. SUSB currently uses OMB Bulletin No. 15-01 for MSA reference. New areas will qualify and the boundaries and/or titles of many existing areas will change. Like counties, these changes are implemented in batch once every 5 years. For questions regarding specific years, please see the MSA Geography Reference page

[<https://web.archive.org/web/20210901053416/https://www.census.gov/programs-surveys/susb/technical-documentation/reference-files.html>] . MSA changes are detailed in the following page: Metropolitan and Micropolitan Guidance for Data Users

[<https://web.archive.org/web/20210901053416/https://www.census.gov/programs-surveys/metro-micro/guidance.html>] .

In 2002 and earlier data, CBP data were published for New England County Metropolitan Areas rather than for MSAs in CT, ME, MA, NH, RI, and VT. Micropolitan areas were first defined in 2003 and no data have been published for prior years.

Treatment of auxiliary establishments

In the 1998 - 2002 SUSB reports, corporate, subsidiary, and regional managing offices were tabulated in NAICS Sector 55. All other auxiliaries were tabulated in NAICS 95. Starting with the 2003 SUSB, corporate, subsidiary, and regional managing offices are still published in NAICS Sector 55, but the other auxiliaries are tabulated in the industry of the service performed. The other auxiliaries were coded into ten separate NAICS classifications listed under the auxiliary section of the 1997 Bridge Between NAICS and SIC.

In 1997 and earlier SUSB data series based on SIC classification, auxiliary establishments were similarly excluded from SIC categories, but were reported as a separate data line for those counties where they were present.

Economic Census

Definitions and coverage differences may affect the direct comparison of Economic Census and SUSB data. See Glossary [</web/20210901053416/https://www.census.gov/programs-surveys/susb/about/glossary.html>] for an explanation of the data items contained in SUSB.

The Economic Census generally uses respondent reported data. SUSB uses a combination of reported Company Organization Survey data and administrative records data. Although efforts are made to resolve significant differences in the data, differences are known to exist. See Methodology [</web/20210901053416/https://www.census.gov/programs-surveys/susb/technical-documentation/methodology.html>] for further information on how the SUSB data are produced.

Some large companies report different activities at the same location as separate profit centers. The SUSB program treats each profit center as a separate establishment. The Economic Census [</web/20210901053416/https://www.census.gov/programs-surveys/economic-census.html>] reporting may combine the profit centers into one establishment. This results in establishment count differences due to differences in how the data are collected.

Definitions

SUSB definitions available in the SUSB Glossary Page

[</web/20210901053416/https://www.census.gov/programs-surveys/susb/about/glossary.html>] .

Last Revised: August 30, 2021